



A Comparison of Maps and Power Spectra Determined from South Pole Telescope and *Planck* Data

Z. Hou^{1,2}, K. Aylor³, B. A. Benson^{1,2,4}, L. E. Bleem^{1,5}, J. E. Carlstrom^{1,2,5,6,7}, C. L. Chang^{1,2,5}, H.-M. Cho⁸, R. Chown⁹, T. M. Crawford^{1,2}, A. T. Crites^{1,2,10}, T. de Haan^{9,11}, M. A. Dobbs^{9,12}, W. B. Everett¹³, B. Follin³, E. M. George^{11,14}, N. W. Halverson^{13,15}, N. L. Harrington¹¹, G. P. Holder^{9,12,16,17}, W. L. Holzapfel¹¹, J. D. Hrubes¹⁸, R. Keisler^{1,6,19}, L. Knox³, A. T. Lee^{11,20}, E. M. Leitch^{1,2}, D. Luong-Van¹⁸, D. P. Marrone²¹, J. J. McMahon²², S. S. Meyer^{1,2,6,7}, M. Millea³, L. M. Mocanu^{1,2}, J. J. Mohr^{23,24,25}, T. Natoli^{1,6,26}, Y. Omori⁹, S. Padin^{1,2}, C. Pryke²⁷, C. L. Reichardt^{11,28}, J. E. Ruhl²⁹, J. T. Sayre^{13,29}, K. K. Schaffer^{1,7,30}, E. Shirokoff^{1,2,11}, Z. Staniszewski^{29,31}, A. A. Stark³², K. T. Story^{1,6,19,33}, K. Vanderlinde^{26,34}, J. D. Vieira^{16,17}, and R. Williamson^{1,2}

¹ Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA; hough@kicp.uchicago.edu

² Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA

³ Department of Physics, University of California, Davis, CA 95616, USA; kmaylor@ucdavis.edu

⁴ Fermi National Accelerator Laboratory, MS209, P.O. Box 500, Batavia, IL 60510, USA

⁵ High Energy Physics Division, Argonne National Laboratory, Argonne, IL 60439, USA

⁶ Department of Physics, University of Chicago, Chicago, IL 60637, USA

⁷ Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA

⁸ SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

⁹ Department of Physics and McGill Space Institute, McGill University, Montreal, Quebec H3A 2T8, Canada

¹⁰ California Institute of Technology, Pasadena, CA 91125, USA

¹¹ Department of Physics, University of California, Berkeley, CA 94720, USA

¹² Canadian Institute for Advanced Research, CIFAR Program in Cosmology and Gravity, Toronto, ON M5G 1Z8, Canada

¹³ Center for Astrophysics and Space Astronomy, Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309, USA

¹⁴ European Southern Observatory, Karl-Schwarzschild-Straße 2, D-85748 Garching, Germany

¹⁵ Department of Physics, University of Colorado, Boulder, CO 80309, USA

¹⁶ Astronomy Department, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA

¹⁷ Department of Physics, University of Illinois Urbana-Champaign, 1110 W. Green Street, Urbana, IL 61801, USA

¹⁸ University of Chicago, Chicago, IL 60637, USA

¹⁹ Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA

²⁰ Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

²¹ Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

²² Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

²³ Faculty of Physics, Ludwig-Maximilians-Universität, D-81679 München, Germany

²⁴ Excellence Cluster Universe, D-85748 Garching, Germany

²⁵ Max-Planck-Institut für extraterrestrische Physik, D-85748 Garching, Germany

²⁶ Dunlap Institute for Astronomy & Astrophysics, University of Toronto, 50 St George St., Toronto, ON M5S 3H4, Canada

²⁷ Department of Physics, University of Minnesota, Minneapolis, MN 55455, USA

²⁸ School of Physics, University of Melbourne, Parkville, VIC 3010, Australia

²⁹ Physics Department, Center for Education and Research in Cosmology and Astrophysics, Case Western Reserve University, Cleveland, OH 44106, USA

³⁰ Liberal Arts Department, School of the Art Institute of Chicago, Chicago, IL 60603, USA

³¹ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

³² Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA

³³ Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA

³⁴ Department of Astronomy & Astrophysics, University of Toronto, 50 St George St., Toronto, ON M5S 3H4, Canada

Received 2017 April 4; revised 2017 December 1; accepted 2017 December 19; published 2018 January 17

Abstract

We study the consistency of 150 GHz data from the South Pole Telescope (SPT) and 143 GHz data from the *Planck* satellite over the patch of sky covered by the SPT-SZ survey. We first visually compare the maps and find that the residuals appear consistent with noise after accounting for differences in angular resolution and filtering. We then calculate (1) the cross-spectrum between two independent halves of SPT data, (2) the cross-spectrum between two independent halves of *Planck* data, and (3) the cross-spectrum between SPT and *Planck* data. We find that the three cross-spectra are well fit ($\text{PTE} = 0.30$) by the null hypothesis in which both experiments have measured the same sky map up to a single free calibration parameter—i.e., we find no evidence for systematic errors in either data set. As a by-product, we improve the precision of the SPT calibration by nearly an order of magnitude, from 2.6% to 0.3% in power. Finally, we compare all three cross-spectra to the full-sky *Planck* power spectrum and find marginal evidence for differences between the power spectra from the SPT-SZ footprint and the full sky. We model these differences as a power law in spherical harmonic multipole number. The best-fit value of this tilt is consistent among the three cross-spectra in the SPT-SZ footprint, implying that the source of this tilt is a sample variance fluctuation in the SPT-SZ region relative to the full sky. The consistency of cosmological parameters derived from these data sets is discussed in a companion paper.

Key words: cosmic background radiation – methods: data analysis

1. Introduction

One of the most remarkable results of modern cosmology is that a simple six-parameter model, usually referred to as the Λ CDM model, can fit the full range of cosmological observations. With the precision of cosmological observables now reaching the level of a few percent, however, several small discrepancies in the inferred parameter values are attracting attention.

These discrepancies show up in three places. The first is in the inferred parameter values from the cosmic microwave background (CMB) compared to some observations of the local universe. For example, the amplitude of local density fluctuations, σ_8 , that is measured from observations of large-scale structure appears lower than the σ_8 value implied by CMB measurements (Planck Collaboration et al. 2016c). There is also some tension between direct measurements of the Hubble parameter H_0 and the value derived from CMB measurements (Riess et al. 2011, 2016; Planck Collaboration et al. 2014c, 2016b).

Second, there are mild (1σ – 2σ) discrepancies between parameter values derived from observations of the CMB by different experiments. In particular, the best-fit parameters of the Λ CDM model given by *Planck* satellite measurements of CMB temperature and polarization power spectra (Planck Collaboration et al. 2016e) are somewhat different from those derived from earlier CMB data, whether from the *Wilkinson Microwave Anisotropy Probe* (WMAP) satellite (Hinshaw et al. 2013), from the combination of WMAP data and data from the South Pole Telescope (SPT; Hou et al. 2014), or from WMAP + SPT plus data from the Atacama Cosmology Telescope (ACT; Calabrese et al. 2013).

Finally, work has been done on the internal consistency of Λ CDM model parameter values from different subsets of the *Planck* data. Addison et al. (2016) recently pointed out that the matter density inferred from *Planck* data at $\ell < 1000$ is 2.5σ discrepant from that inferred from *Planck* data at $\ell > 1000$. In response, Planck Collaboration et al. (2017) show that after correcting for certain approximations in the Addison et al. (2016) analysis, and taking into account the fact that matter density had been singled out as the most discrepant parameter, the global discrepancy is only 1.6σ .

Taken together, these low-level discrepancies have led some to speculate that we are seeing evidence of a potential failure of the Λ CDM model (Battye & Moss 2014; Wyman et al. 2014), systematic errors in the analysis of the low-redshift probes (Efstathiou et al. 2014), or systematic errors in the *Planck* data (Planck Collaboration et al. 2017). The analysis presented in this paper is motivated by the fact that it is the high- ℓ temperature data from *Planck* that are driving the parameter shifts of interest (Addison et al. 2016; Planck Collaboration et al. 2017). In terms of the measurement uncertainties, data from the 2540 deg^2 SPT-Sunyaev–Zeldovich (SZ) survey yield constraints within a factor of 2 of the *Planck* constraints at $\ell \gtrsim 1700$ and better than *Planck* at $\ell \gtrsim 2100$ (Story et al. 2013, hereafter S13). The SPT-SZ data are thus a logical choice for consistency checks of the high- ℓ *Planck* data.

In this work, we present the first comparison between *Planck* and SPT data over the same patch of sky. Checks have previously been performed at the power spectrum and cosmological-parameter level, for instance, in Planck Collaboration et al. (2014c, 2016b), and the two data sets have been shown to be roughly consistent, but again with some low-level discrepancies.

The strength of power spectrum and parameter comparisons using the full data sets is limited by the sample variance of the SPT data at lower ℓ and *Planck* noise at higher ℓ . By limiting the comparison to CMB modes measured by both experiments, we can greatly reduce the sample variance to sharpen the consistency tests between the two data sets.

Here we compare the SPT-SZ data in the 150 GHz band and *Planck* full-mission data in the 143 GHz band, restricted to the SPT-SZ observing region. We calculate angular cross-spectra of the SPT and *Planck* maps (150×143) and, for comparison, the cross-spectrum of one half of SPT data with the other half (150×150) and the cross-spectrum of one half of *Planck* data with the other half (143×143). We calculate the difference between the SPT 150×150 spectrum and the other two, as well as the ratio of the SPT 150×150 spectrum to the others. Using simulated observations of mock skies (including realistic noise for both experiments), we calculate the expected uncertainty in these differences and ratios, and we use this to calculate the χ^2 and probability to exceed (PTE) this χ^2 under the null hypothesis that there are no systematic biases between the experiments. This investigation is similar to that performed by Louis et al. (2014) on ACT and *Planck* data over 592 deg^2 of sky and two observing bands (143/148 GHz and 217/218 GHz), the conclusion of which was that ACT and *Planck* measured statistically consistent CMB fluctuations over that patch of sky.

This paper is organized as follows. In Section 2 we discuss the SPT 150 GHz and *Planck* 143 GHz temperature maps in the 2540 deg^2 patch of sky that constitutes the whole of the SPT-SZ survey region. In Section 3 we present how power spectra are calculated from the maps and the simulations generated to de-bias the data and build the covariance matrix. In Section 4 we compare the power spectra and use simulations to test the null hypothesis that the two experiments are measuring the same sky, and we discuss these results in the context of comparisons between SPT and full-sky *Planck* data. We present our conclusions in Section 5.

2. Data

The main goal of this work is to compare maps and power spectra from the SPT 150 GHz and *Planck* 143 GHz data sets within the 2540 deg^2 SPT-SZ survey area. As an overview of the two data sets, Figure 1 shows a sky map of *Planck* HFI 143 GHz full-mission data, with the SPT-SZ survey area outlined by a solid black curve. In this section, we discuss details of the mapmaking and instrumental characteristics of each experiment.

2.1. SPT

The SPT is a 10 m telescope located at the Amundsen–Scott South Pole station. From 2008 to 2011, the first camera on the SPT, a three-band bolometer array known as the SPT-SZ camera, was used to conduct a survey of $\sim 2500\text{ deg}^2$ of the southern sky with low Galactic dust contamination, referred to as the SPT-SZ survey. As shown in Figure 1, the survey area is a contiguous region extending from 20^{h} to 7^{h} in right ascension (R.A.) and from -65° to -40° in declination. The survey was conducted by observing a series of 19 areas, ranging in area from roughly 100 to 300 deg^2 , which together form the full survey region (see, e.g., S13, Figure 2). If not specified, in this

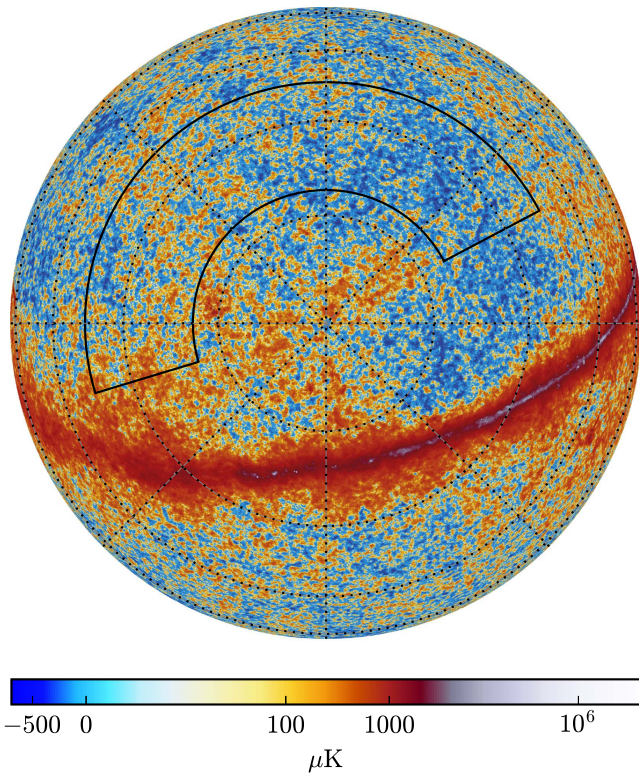


Figure 1. Celestial southern hemisphere of CMB data from *Planck* HFI 143 GHz data. The black curve outlines the SPT-SZ survey coverage. The image is oriented such that the line of 0^h right ascension extends from the center to the top of the hemisphere and right ascension increases counterclockwise.

paper we use “field” to refer to individual observing areas of SPT-SZ.

There were approximately 200 observations of each field, with each individual observation taking roughly 2 hr. Estimates of the primary CMB temperature power spectrum from this survey, as well as the resulting cosmological interpretation, are discussed in Keisler et al. (2011, hereafter K11), S13, and Hou et al. (2014). Our analysis is identical to that of S13 up to the single-observation map step. We briefly review that part of the analysis here and refer the reader to K11 and S13 for more details.

To process time-ordered data (TOD) into maps, the TOD from each SPT detector are first filtered and multiplied by a calibration factor. The filtering steps important for this analysis are a high-pass filter and the subtraction of a common mode across all detectors in each of the six 160-element modules in the focal plane. The high-pass filter cuts off signals below a temporal frequency corresponding to an angular frequency along the scan direction of $\ell = 270$. The common-mode subtraction acts as an isotropic high-pass filter with a cutoff at roughly $\ell = 500$. In both of these filtering steps, bright point sources ($S > 50$ mJy at 150 GHz) are masked to avoid large filtering artifacts.

The TOD from individual SPT detectors are then binned into maps using inverse-variance weighting, i.e., TOD samples corresponding to times when an individual detector was pointed at a particular pixel are averaged together using inverse-variance weighting and then assigned to that pixel. Maps are made using the oblique Lambert equal-area azimuthal projection (Snyder 1987) with 1 arcmin square pixels.

The SPT maps are simply binned and averaged maps of filtered data and thus are biased representations of the sky. The signal in these maps is the true sky signal convolved with the instrument beam, the effect of the TOD filtering, and the effect of binning. Beams are discussed below, and the filter transfer function is estimated through simulations, as discussed in Section 3.3.

We use the S13 estimates of the SPT 150 GHz beam transfer function B_ℓ and its uncertainty, and we refer the reader to S13 for more details. Briefly, the main lobe is measured using bright point sources in the survey fields, while the sidelobes are measured using observations of Jupiter. Venus observations are used to stitch the two together. The main lobe of the beam is well approximated by a 1.2 arcmin FWHM Gaussian. The beam uncertainty arises from several statistical and systematic effects, including residual atmospheric noise in the maps of Venus and Jupiter, and the weak dependence of B_ℓ on the choice of radius used to stitch the inner and outer beam maps.

The TOD from each detector in each observation is calibrated using the response to an internal thermal source, which is in turn tied to the brightness of the Galactic H II region RCW38. For details of this calibration, see Schaffer et al. (2011). The full-depth maps are then compared to CMB satellite data to provide the overall absolute calibration. In K11 and S13, the power spectrum of the full-depth maps was compared to the full-sky *WMAP* estimate of the CMB power spectrum in the multipole range $650 \leq \ell \leq 1000$. For this work, we initially use the calibration determined in George et al. (2015), using the comparison of SPT power spectra to the full-sky *Planck* power spectrum in the multipole range $670 \leq \ell \leq 1170$; however, the map-based comparison to *Planck* undertaken here ends up providing a significantly more precise absolute calibration, as detailed in Section 4.2.

2.2. Planck

The *Planck* satellite (Planck Collaboration et al. 2014a) was launched in 2009 by the European Space Agency, with the goal of measuring the CMB temperature and polarization anisotropy with significantly better sensitivity, angular resolution, and wavelength coverage than was achieved by its space predecessor, the *WMAP* mission (Bennett et al. 2013). *Planck* mapped the full sky in nine bands, ranging in frequency from 30 to 857 GHz. In this work, we use HFI data from the 2015 data release. Specifically, we use the 143 GHz full-mission, halfmission-1, and halfmission-2 maps downloaded from the NASA/IPAC Infrared Science Archive (IRSA).³⁵

In this work, we use a cross-spectrum pipeline similar to that used in K11 and S13 to compare SPT and *Planck* data on the SPT-SZ sky patch. To use this pipeline for *Planck* data, the *Planck* maps must have the same pixelization, projection, and field definitions as the SPT maps. To achieve this, we create mock TOD using the *Planck* 143 GHz full-mission or half-mission maps and the SPT pointing and detector weight information from individual SPT observations. The mock TOD are then binned into maps in the same manner as the real SPT TOD, using the same projection and pixelization. Instead of using the SPT bandpass filtering, a much simpler high-pass filter is applied to the *Planck* mock TOD to simply remove the signals on scales with $\ell < 100$. Our simulations show that applying this simple high-pass filter greatly improves the

³⁵ http://irsa.ipac.caltech.edu/data/Planck/release_2/all-sky-maps/

numerical stability of our unbiased power spectrum calculations. In the power spectrum analysis described below, we use the projected *Planck* maps generated from the mapmaking pipeline with this simple high-pass filter applied to the *Planck* mock TOD.

The *Planck* 143 GHz beam has been measured using the *Planck* TOD around planets, such as Mars, Jupiter, and Saturn (Planck Collaboration et al. 2014b, 2016f). The beam window function $B(\ell)$ for individual frequency detector sets has been released by the Planck Collaboration as part of the “Reduced Instrument Model” (RIMO).³⁶ From the 2013 *Planck* data release to the 2015 data release, there was a marked improvement in the beam characterization, such that the uncertainty on the 2015 beam window functions for 143 GHz data is at the 0.1% level for the ℓ range of interest to this work.

3. Power Spectrum

We now turn to the methodology for calculating unbiased cross-spectra of SPT 150 GHz maps and *Planck* 143 GHz maps. We introduce the power spectrum estimator in Section 3.1. Simulations are used in several places while estimating the power spectra; we describe how these simulations are generated in Section 3.2. We discuss how these simulations are used to estimate the SPT filter transfer function in Section 3.3 and to calculate the noise and sample variance parts of the power spectrum covariance in Section 3.4.1. We describe how beam uncertainty is incorporated into the analysis in Section 3.4.2. Finally, we discuss how to calculate the window functions needed to compare the binned cross-spectrum estimates—or “bandpowers”—to a theory spectrum in Section 3.4.3.

3.1. Power Spectrum Estimator

We use an estimator similar to that used in K11 and S13 to calculate the various cross-spectra in this work. Each cross-spectrum is calculated by correlating maps made from different observations of the same field—either correlating full-depth SPT maps with full-depth *Planck* maps or correlating two half-depth maps from the same experiment. In the latter case, there is a noise penalty relative to K11 or S13, who used $\mathcal{O}(100)$ independent maps depending on the field. However, the noise penalty is largely insignificant for the SPT maps (less than 2% in power) on the angular scales of interest, and it is unavoidable for the *Planck* maps because a larger number of independent splits are not publicly available. We mask and zero-pad each map before calculating its two-dimensional Fourier transform, \tilde{m}_ℓ .³⁷ The mask is a product of an apodization window and a point-source mask.³⁸ The maps and masks are zero-padded before Fourier-transforming such that each Fourier-space pixel has a width of $\delta_\ell = 5$. The raw bandpowers are the binned average of the cross-spectra between two maps A and B within

a multipole bin b :

$$\hat{D}_b^{A \times B} = \left\langle \frac{\ell(\ell+1)}{2\pi} H_\ell \text{Re} [\tilde{m}_\ell^A \tilde{m}_\ell^{B*}] \right\rangle_{\ell \in b}. \quad (1)$$

For the SPT-only power spectrum, A and B are the two SPT half-survey maps; for the SPT and *Planck* cross-spectrum, A and B are the full-mission maps of each experiments; and for the *Planck*-only 143 power spectrum, A and B are the two *Planck* half-mission maps. H_ℓ in the above equation represents the two-dimensional weighting of the Fourier modes that was used by S13 to handle the anisotropic noise in the SPT maps. While this weighting is suboptimal for *Planck* data, we still use the same S13-derived weighting for all bandpowers to minimize the differential sample variance between cross-spectra.

Since the input maps are biased estimates of the true sky, due to effects such as the application of a mask to the maps, the raw bandpowers \hat{D}_b are a biased estimate of the true bandpowers, D_b . The biased and unbiased estimates are related by

$$\hat{D}_{bb'} \equiv K_{bb'} D_b, \quad (2)$$

where

$$K_{bb'} = P_{b\ell'} (M_{\ell\ell'} F_{\ell'} B_{\ell'}^2) Q_{\ell'b'}, \quad (3)$$

$Q_{\ell'b'}$ is the binning operator and P_ℓ is its reciprocal, $M_{\ell\ell'}$ is the mode-coupling matrix, F_ℓ is the filter transfer function that accounts for the signal suppressed by TOD filtering, and B_ℓ is the beam function. For details on the unbiasing procedure, see K11, S13, and Hivon et al. (2002).

The unbiased bandpowers are calculated on a field-by-field basis and then combined. We combine the bandpowers obtained from individual fields using the effective area of single fields as the weighting, as in K11 and S13.

3.2. Simulations

We use simulations both in the calculation of the bandpowers and to characterize the degree of consistency between the SPT and *Planck* cross-spectra over the SPT-SZ survey region. In this section, we turn our attention to how these simulations are created and used. The final product of this procedure is 400 sets of simulated unbiased bandpowers for each of the three combinations of data (150×150 , 150×143 , and 143×143).

3.2.1. Sky Signals

Our simulations include the following components: (1) a gravitationally lensed CMB signal, (2) thermal and kinematic Sunyaev–Zel’dovich (SZ) signals, (3) the cosmic infrared background (CIB) signal, and (4) emission from radio galaxies. We generate 400 realizations of the lensed CMB with LensPix (Lewis 2005), based on the best-fit Λ CDM model from *Planck* 2015 TT, TE, EE + lowP + lensing (Planck Collaboration et al. 2016b). Modes are generated out to $\ell_{\text{max}} = 8000$, well above the angular multipoles, $\ell < 2500$, used in this comparison. The maps are stored in HEALPix (Górski et al. 2005) format with resolution parameter $N_{\text{side}} = 8192$.

Unlike S13, which used Gaussian realizations for all extragalactic foregrounds, here we generate realizations of individual sources for the bright CIB galaxies ($6.4 \text{ mJy} < S < 50 \text{ mJy}$) and all radio galaxies (up to the flux

³⁶ http://irsa.ipac.caltech.edu/data/Planck/release_2/ancillary-data/

³⁷ When calculating angular power spectra in this work, we use the flat-sky approximation, in which we replace spherical harmonic transforms with two-dimensional Fourier transforms.

³⁸ In this work, we use the same point-source masks used in S13 but slightly different apodization masks; we have confirmed using simulations that any effect of the different masking on power spectra or cosmological parameters is negligible for this analysis.

cut at 50 mJy). The bright CIB galaxies are drawn from the modeled dN/dS of Cai et al. (2013), while the radio galaxy dN/dS is taken from De Zotti et al. (2005). In both cases, the amplitudes of the dN/dS distribution are calibrated by the actual observations of Mocanu et al. (2013) at 150 GHz.

The other extragalactic foregrounds, the thermal and kinematic SZ signals and the low-flux CIB, are treated as Gaussian. The shapes of the thermal and kinematic SZ angular power spectra are taken from the Shaw et al. (2010) and Shaw et al. (2012) models, respectively, with the amplitudes set to the median values from George et al. (2015), $D_{3000}^{\text{SZ}} = 4.38 \mu\text{K}^2$ and $D_{3000}^{\text{KSZ}} = 1.57 \mu\text{K}^2$. We similarly draw upon the median values and templates from George et al. (2015) for the CIB terms. The clustered CIB spectrum is taken to follow $D_\ell \propto \ell^{0.8}$ with an amplitude $D_{3000}^{\text{C}} = 3.46 \mu\text{K}^2$. The shot-noise or “Poisson” CIB power from galaxies dimmer than 6.4 mJy is taken to be $D_{3000}^{\text{P}} = 9.16 \mu\text{K}^2$.

Each component is scaled appropriately to the effective frequency of the SPT or *Planck* maps and then co-added together to create the final sky realization.

3.2.2. Planck Noise Simulations

The Planck Collaboration released the 8th Full Focal Plane (FFP8) simulation in the 2015 data release (Planck Collaboration et al. 2016d). There are 1000 full-mission FFP8 noise simulations available at NERSC.³⁹ We use the FFP8 full-mission simulations to create noisy *Planck*-like realizations to characterize the noise contribution to the SPT-*Planck* cross (150×143) bandpowers as discussed in Section 4.2.

The FFP8 noise simulations are only available for the full-mission data and thus cannot be used for the *Planck* half-mission cross-spectrum. To create unbiased *Planck* 143 GHz bandpowers within the SPT area, we also generate half-mission noise simulations to characterize the noise contribution to the *Planck-Planck* (143×143) bandpowers as presented in Section 4.2. The half-mission noise simulations are based on the pixel-space noise variance released by the Planck Collaboration along with the half-mission observation maps. We approximate the *Planck* map noise as Gaussian and uncorrelated between pixels (“white”) for the half-mission noise simulations.

To test this white-noise approximation, we compare the noise power spectrum over the SPT survey area from 100 of these Gaussian white-noise realizations to 100 FFP8 full-mission noise simulations. The average power ratio is plotted in Figure 2. Clearly, the white-noise realizations overestimate the noise power for $150 < \ell < 2200$, i.e., most angular multipoles of interest. This implies that the *Planck* noise contribution is overestimated in our *Planck*-only bandpowers, at a maximum level of roughly 15% and a mean level of $\lesssim 10\%$ across the ℓ range of interest. We discuss the possible impact of this overestimate on our results in Section 4.2 and find that our main conclusions would likely be unchanged had we used more realistic noise simulations for the *Planck*-only bandpowers.

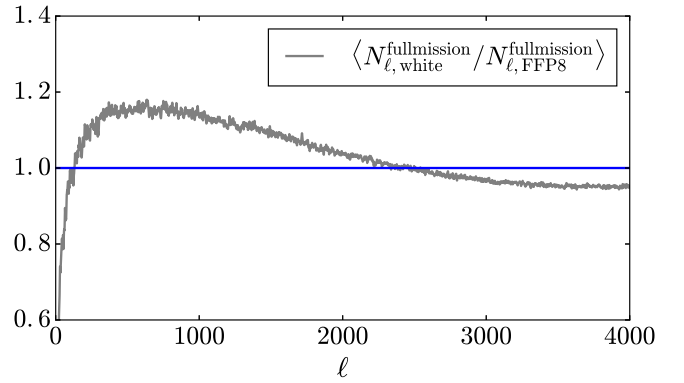


Figure 2. Averaged ratio of the noise power spectrum between the white-noise realizations and *Planck* FFP8 noise simulations within the SPT survey area. While we use the full FFP8 noise simulations for the most constraining set of bandpowers (SPT \times *Planck*), we use the simpler white-noise realizations when estimating uncertainties for the cross-spectrum between half-survey *Planck* maps.

3.2.3. SPT Noise Realizations

SPT noise realizations are created directly from the data. For each individual SPT observation of each SPT-SZ survey field, we create a residual map that is the difference between a map made from all left-going telescope scans and a map made from all right-going telescope scans. There are approximately 200 jackknife maps for each field (the number of individual ~ 2 hr observations). We use these maps to create the noise part of the simulated SPT observations. For each field, we multiply $+1$ or -1 randomly to the residual maps and co-add them together to form one noise realization. Using this method, we create 400 noise realizations.

With the same method, we also create SPT noise realizations of the first and second half sets of observations by co-adding the residual maps of half of the observations within one field. Similar to the half-mission noise simulations for *Planck*, we use the SPT noise simulations to characterize the noise contribution to the SPT-only bandpowers.

3.3. Filtering Transfer Function

The noise-free simulated observations are used to calculate F_ℓ , the filtering transfer function, and H_ℓ , the two-dimensional weight in the power spectrum calculation. Given the input power spectrum of the simulations, the effective transfer function can be derived by comparing the power spectrum of the simulations with the known input spectrum using an iterative scheme (Hivon et al. 2002). This method was used in K11 and S13; in this analysis, we make adjustments to Equations (6) and (7) in Keisler et al. (2011) to include the *Planck* beam and pixel window function for the SPT-*Planck* cross-spectrum and *Planck*-only spectrum.

3.4. Consistency Metrics

We have three sets of unbiased bandpowers derived from SPT and *Planck* data, $\mathcal{D}_b^{150 \times 150}$, $\mathcal{D}_b^{150 \times 143}$, and $\mathcal{D}_b^{143 \times 143}$. In this analysis, we use the difference between these bandpowers, in particular the residual and the ratio between the bandpower sets, to quantitatively characterize the consistency between the SPT and *Planck* data sets. This section presents the metrics we use for these consistency tests.

³⁹ <https://crd.lbl.gov/departments/computational-science/c3/c3-research/cosmic-microwave-background/cmb-data-at-nersc/>

3.4.1. Covariance Estimation

We estimate the bandpower sample variance and noise variance using 400 sets of simulated signal+noise maps. The final bandpower covariance matrix will also include a contribution due to beam uncertainties; the calculation of the beam covariance is detailed in the next section. For a single field and data combination, the bandpower covariance, $\Xi^{X \times Y}$, is calculated simply as

$$\Xi^{X \times Y} = \langle (\mathbf{D}_{b,\text{sim}}^{X \times Y} - \bar{\mathbf{D}}_{b,\text{sim}}^{X \times Y})(\mathbf{D}_{b',\text{sim}}^{X \times Y} - \bar{\mathbf{D}}_{b',\text{sim}}^{X \times Y}) \rangle, \quad (4)$$

where $\bar{\mathbf{D}}_{b,\text{sim}}^{X \times Y}$ is the mean over all 400 simulations for cross-bandpowers $X \times Y \in [150 \times 150, 150 \times 143, 143 \times 143]$. The simulated 150×143 cross-bandpowers $\mathbf{D}_{b,\text{sim}}^{150 \times 143}$ are derived from cross-spectra of simulated SPT maps including realizations of SPT full-observation noise and simulated *Planck* maps of the same underlying sky signal including *Planck* FFP8 noise simulations. For $\mathbf{D}_{b,\text{sim}}^{150 \times 150}$, the simulated bandpowers are calculated by the cross-correlation of the two sets of simulated SPT maps with half-observation noise realizations. Similarly, the simulated $\mathbf{D}_{b,\text{sim}}^{143 \times 143}$ bandpowers are obtained using two sets of simulated *Planck* maps including *Planck* half-mission white-noise realizations. The full bandpower covariance matrix is then obtained by combining the individual-field covariance matrices with the square of the area weighting used to combine the bandpowers themselves.

As we will evaluate consistency by looking at the differences and ratios among sets of bandpowers, we also need the covariance matrix for these quantities. For the differences or residual bandpowers, $\Delta \mathbf{D}_b \equiv \mathbf{D}_b^{i \times j} - \mathbf{D}_b^{m \times n}$ (where $i, j, m, n \in \{150, 143\}$), the noise-plus-sample-variance part of the covariance is easily estimated from the 400 simulations:

$$\Xi_{\text{resid}} = \langle \Delta \mathbf{D}_{b,\text{sim}} \Delta \mathbf{D}_{b',\text{sim}} \rangle. \quad (5)$$

For the ratios relative to the 150×150 bandpowers, the covariance can be expressed as

$$\Xi_{\text{ratio}} = \left\langle \left(\frac{\mathbf{D}_{b,\text{sim}}^{i \times j}}{\mathbf{D}_{b,\text{sim}}^{150 \times 150}} - 1 \right) \left(\frac{\mathbf{D}_{b',\text{sim}}^{i \times j}}{\mathbf{D}_{b',\text{sim}}^{150 \times 150}} - 1 \right) \right\rangle. \quad (6)$$

3.4.2. Beam Uncertainty

In this section, we present how beam uncertainties are handled for the bandpower comparison. Beam uncertainties appear as a second covariance term, in addition to the noise-plus-sample-variance term presented above, because the simulations did not include beam uncertainties.

We begin by estimating the eigenmodes of the *Planck* and SPT beam covariance matrices. The *Planck* beam eigenmodes can be obtained from RIMO in the 2015 data release, and the SPT beam eigenmodes can be derived from the beam correlation matrix calculated in S13. For each eigenmode, we can calculate the fractional beam uncertainty as a function of multipole, $\delta b_\ell / b_\ell$, and propagate this linearly to the bandpower space according to

$$\delta D_{b,\text{beam}}^{i \times j} = - \sum_{\ell} W_{b\ell}^{i \times j} D_{\ell}^{i \times j, \text{fid}} \left(\frac{\delta b_{\ell}^i}{b_{\ell}^i} + \frac{\delta b_{\ell}^j}{b_{\ell}^j} \right), \quad (7)$$

where $i, j \in \{150, 143\}$, $W_{b\ell}^{i \times j}$ is the bandpower window function, and $D_{\ell}^{i \times j, \text{fid}}$ is the fiducial power spectrum including

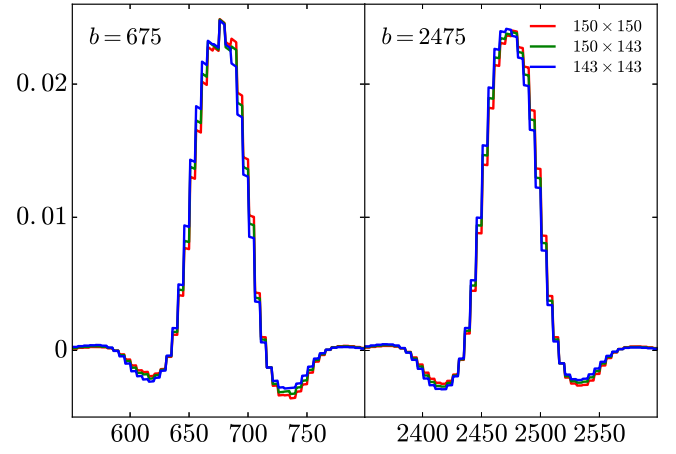


Figure 3. Window functions $W_{b\ell}$ for the $D_b^{150 \times 150}$ (red), $D_b^{150 \times 143}$ (green), and $D_b^{143 \times 143}$ (blue) bandpowers, for bins centered at $\ell = 675$ and 2475 . In both panels, $W_{b\ell}^{143 \times 143}$ has a higher (lower) weight on ℓ s lower (higher) than the bin center of $W_{b\ell}^{150 \times 150}$, and $W_{b\ell}^{150 \times 143}$ sits between the two. Due to the shape of the CMB power spectrum, D_{ℓ} , such difference leads to an ℓ -dependent bias when directly comparing the three sets of bandpowers.

CMB and extragalactic foregrounds for that frequency combination.

For the consistency tests, we need the beam covariance for bandpower differences and bandpower ratios. For the bandpower differences, the above equation leads straightforwardly to

$$\Xi_{bb',\text{beam}}^{(i \times j)(m \times n)} = \sum_{\text{e-modes}} (\delta D_{b,\text{beam}}^{i \times j} - \delta D_{b,\text{beam}}^{150 \times 150}) \times (\delta D_{b',\text{beam}}^{m \times n} - \delta D_{b',\text{beam}}^{150 \times 150}), \quad (8)$$

where $i, j, m, n \in \{150, 143\}$ and the sum is taken over all eigenmodes of the beam covariance matrices.

For the bandpower ratios, the beam covariance can be written as

$$\Xi_{bb',\text{beam}}^{(i \times j)(m \times n)} = \sum_{\text{e-modes}} \left(\frac{\delta D_{b,\text{beam}}^{i \times j} - \delta D_{b,\text{beam}}^{150 \times 150}}{D_b^{150 \times 150}} \right) \times \left(\frac{\delta D_{b',\text{beam}}^{m \times n} - \delta D_{b',\text{beam}}^{150 \times 150}}{D_{b'}^{150 \times 150}} \right). \quad (9)$$

The total bandpower covariance for a consistency test is then the sum of the noise-plus-sample-variance term from the previous section and the appropriate beam covariance term above.

3.4.3. Bandpower Window Function Correction

To characterize the level of consistency between the three sets of bandpowers (150×150 , 150×143 , and 143×143), we choose one set, $D_b^{150 \times 150}$, as the fiducial that is subtracted from or divided into the other two sets (see Section 4.2). However, the residual bandpowers and bandpower ratios are biased owing to the difference of bandpower window functions (e.g., Knox 1999). Due to the different filtering and the beam transfer functions of the two experiments, the bandpower window functions are different between bandpower sets. In Figure 3, the window functions of the three sets of bandpowers are illustrated for two bins with effective centers at $\ell = 675$ and $\ell = 2475$. In both of these bins, relative to the 150×150 window functions, the 143×143 and 150×143 bandpowers

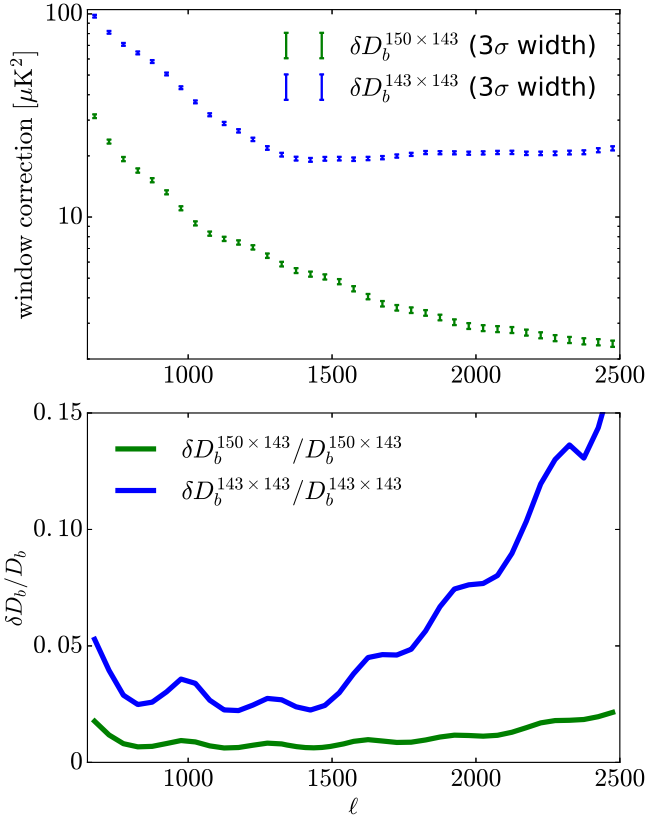


Figure 4. Top panel: bandpower window function correction as given by Equation (10) for 150×143 (green) and 143×143 (blue), with the fiducial D_ℓ including the *Planck* 2015 best-fit cosmology and the best-fit foreground models. The error bars indicate the 3σ variation in the window function corrections marginalized over a Markov chain based on *WMAP*9+SPT data. Bottom panel: ratio between the window function correction δD_b^i and the corresponding bandpower D_b^i for $i = 150 \times 143$ (green) and 143×143 (blue).

receive more weight from multipoles lower than the bin center and less from multipoles higher than the bin center. These differences, if left unaddressed, lead to a bias in the bandpower differences that is dependent, to some degree, on the assumed cosmological model.

Before comparing the bandpower sets, we need to correct for this bias. The correction is calculated as follows:

$$\delta D_b^i = \sum_{\ell} (W_{\ell}^i - W_{\ell}^{150 \times 150}) D_{\ell}^{\text{fid}}, \quad (10)$$

where $i = 150 \times 143$, or 143×143 . The *Planck* 2015 best-fit cosmology plus the best-fit extragalactic foregrounds are used as the fiducial model. We subtract this correction term from the 150×143 and 143×143 bandpowers to remove the bias caused by the bandpower window function differences. In Figure 4, the top panel shows the shape of the correction δD_b and the bottom panel shows the ratio between δD_b and the corresponding D_b . For 150×143 , δD_b yields a roughly flat 1%–2% correction to the bandpower, while for 143×143 , the correction is more important at higher multipoles, up to $\sim 15\%$ at $\ell = 2500$. The window function correction is clearly critical for the consistency analysis between these three sets of bandpowers.

We now investigate the model dependence of the window function corrections. To characterize this dependence, we calculate the window function corrections for a distribution of

input D_{ℓ} . The distribution is obtained by randomly sampling the cosmology and foreground models from a Markov chain based on *WMAP*9+SPT data. The variation of δD_b in Equation (10) is illustrated in the top panel of Figure 4, with the error bars indicating the 3σ variation of the corrections. Within the band range of interest, this variation level is less than 1% of the bandpower standard error; therefore, we ignore this very small uncertainty in our analysis.

4. Results

In this section, we present the results of the comparison between SPT 150 GHz data and *Planck* 143 GHz data. We first present a qualitative map-level comparison of the two data sets in a common sky area. We then present the major result of the paper: the bandpower comparison among three sets of bandpowers calculated on the 2540 deg² SPT-SZ sky patch: $\mathcal{D}_b^{150 \times 150}$, $\mathcal{D}_b^{150 \times 143}$, and $\mathcal{D}_b^{143 \times 143}$. Setting $\mathcal{D}_b^{150 \times 150}$ as the fiducial and subtracting it from the other two, we use the residual as the quantity to characterize the level of consistency between SPT 150 GHz and *Planck* 143 GHz data. As a by-product, for the first time we recalibrate the SPT data to the *Planck* data on the same patch of sky.

We also characterize the ratios between these sets of bandpowers. Note that the residual and ratio tests are not intended to be independent checks, as they contain nearly the same information presented in different ways. The reason we use both metrics is that the systematic errors most likely to cause differences between SPT and *Planck* data (such as unmodeled foreground residuals or beam systematics) fall into two broad categories: additive or multiplicative systematics. Additive systematics will show up most obviously in the residual bandpower test, while multiplicative systematics will show up most obviously in a bandpower ratio test.

Finally, we compare each set of 2540 deg² bandpowers to the full-sky *Planck* 143 GHz bandpowers. Obviously such bandpower comparisons are not distinct from the map comparison. Rather, the residuals and the ratios of these sets of bandpowers are the quantitative version of the by-eye map comparison.

4.1. Map-level Comparison

The SPT and *Planck* maps within the same patch are presented in Figure 5 for visual comparison. The top panels show the filtered SPT map described in Section 2.1 on the left and the projected *Planck* full-mission 143 GHz map on the right. Some bright point sources can be identified by eye in both the SPT and *Planck* maps, but the maps do not resemble each other, because the *Planck* map is dominated by the degree-scale CMB anisotropy filtered out of the SPT data, and the SPT map shows more small-scale structure owing to its higher angular resolution. In the bottom panels of Figure 5, we show SPT and *Planck* maps of the same spatial modes. The bottom left panel shows the SPT map from the top left panel convolved with the difference between the SPT 150 GHz and *Planck* 143 GHz beams; the bottom right panel shows a map made by observing a *Planck* sky with the SPT 150 GHz scan strategy and TOD filtering. Though the *Planck* map has a visibly higher noise level (as expected), the signals in the two maps appear nearly identical. Figure 6 shows the difference between the bottom left and bottom right maps from Figure 5, on the same color scale, along with a simulated difference map

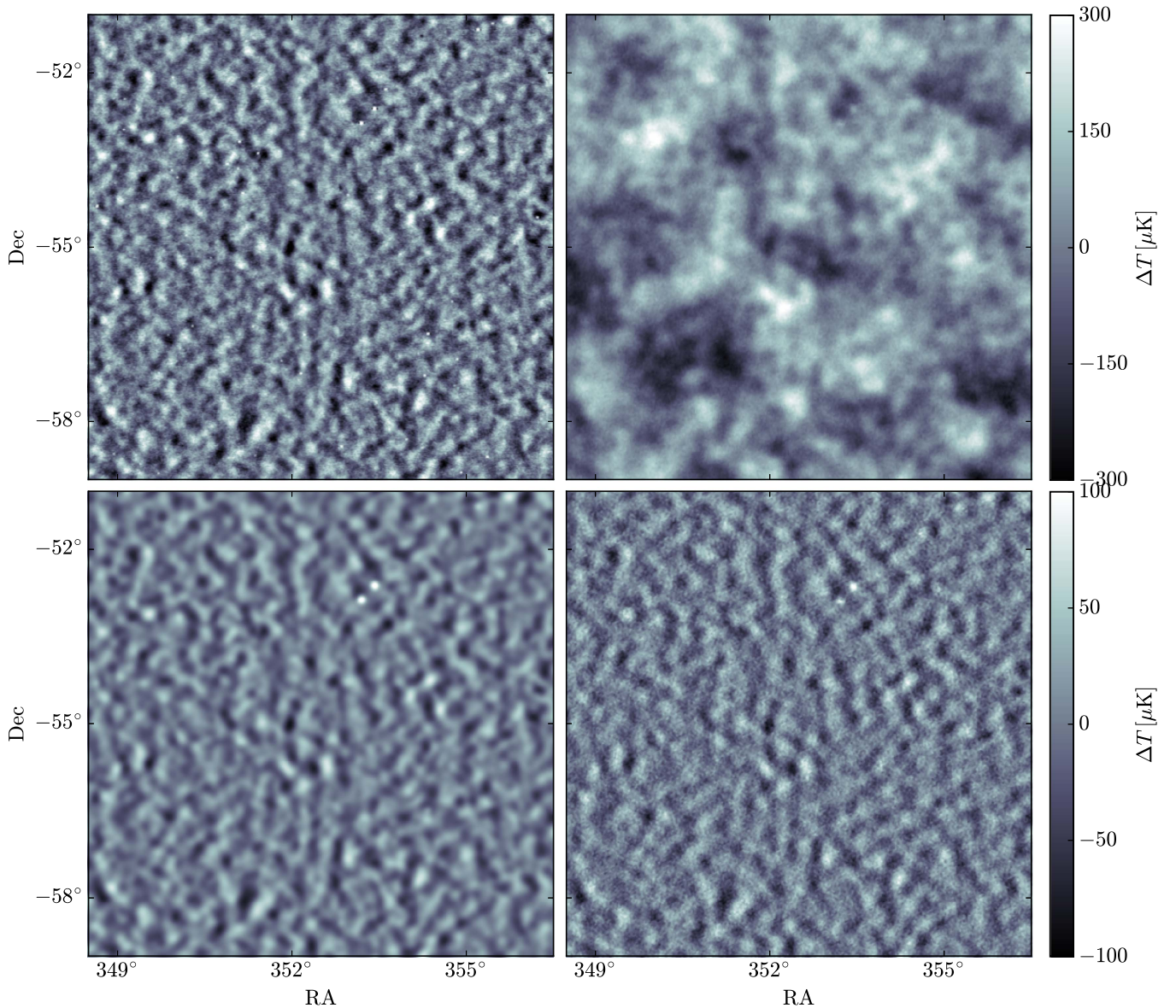


Figure 5. Temperature maps of a portion of one SPT-SZ survey field. Top left: SPT-SZ 150 GHz map of this region. Modes with $\ell \lesssim 600$ are strongly suppressed in this map by the high-pass filter applied to the time-ordered data. Top right: *Planck* 143 GHz full-mission map of the same area. Bottom left: SPT map from the top left panel smoothed to have the same resolution as the *Planck* map. Bottom right: *Planck* map from the top right panel with the SPT-SZ high-pass filter applied. The difference between the bottom left and bottom right panels is shown in Figure 6. Note: the grayscale range of the top right panel is $[-300 \mu\text{K}, 300 \mu\text{K}]$, and the grayscale range of the other three panels is $[-100 \mu\text{K}, 100 \mu\text{K}]$.

for comparison. The feature at the location of one of the bright point sources is most likely caused by temporal variability of the source, as the SPT and *Planck* data were not taken simultaneously.

To make this result more quantitative, we could match the beam and filtering between these two data sets over the full SPT-SZ survey region, mask point sources, calculate the power in the difference map, and compare that power to the expected power from noise alone. In the next section we do a nearly equivalent but somewhat simpler calculation. Using the cross-spectrum formalism outlined in Section 3, we calculate the *Planck* 143×143 power spectrum, the SPT 150×150 power spectrum, and the SPT-*Planck* 150×143 cross-spectrum, and we calculate the χ^2 of the null hypothesis that these three spectra are measuring the same power. This calculation does not completely eliminate sample variance, as the perfectly

mode-matched difference-map power spectrum would, but it strongly reduces it (see Figure 8 below).

4.2. Bandpower Comparison and Recalibration

In this section, we present the comparison between the three sets of window-function-corrected bandpowers, which we denote as $\mathcal{D}_b^{150 \times 150}$, $\mathcal{D}_b^{150 \times 143}$, and $\mathcal{D}_b^{143 \times 143}$. We first show all three sets of bandpowers in the top panel of Figure 7. The error bars contain contributions from sample variance, noise variance, and beam uncertainties. By eye, the three sets of bandpowers look very consistent. At low ℓ , the scatter among the three sets of points is much smaller than the errors on any one set; this is because the errors on all three are dominated by sample variance in this regime, and sample variance is highly correlated among the three data sets. This is also why the three sets of error bars are nearly the same size at low ℓ (though the

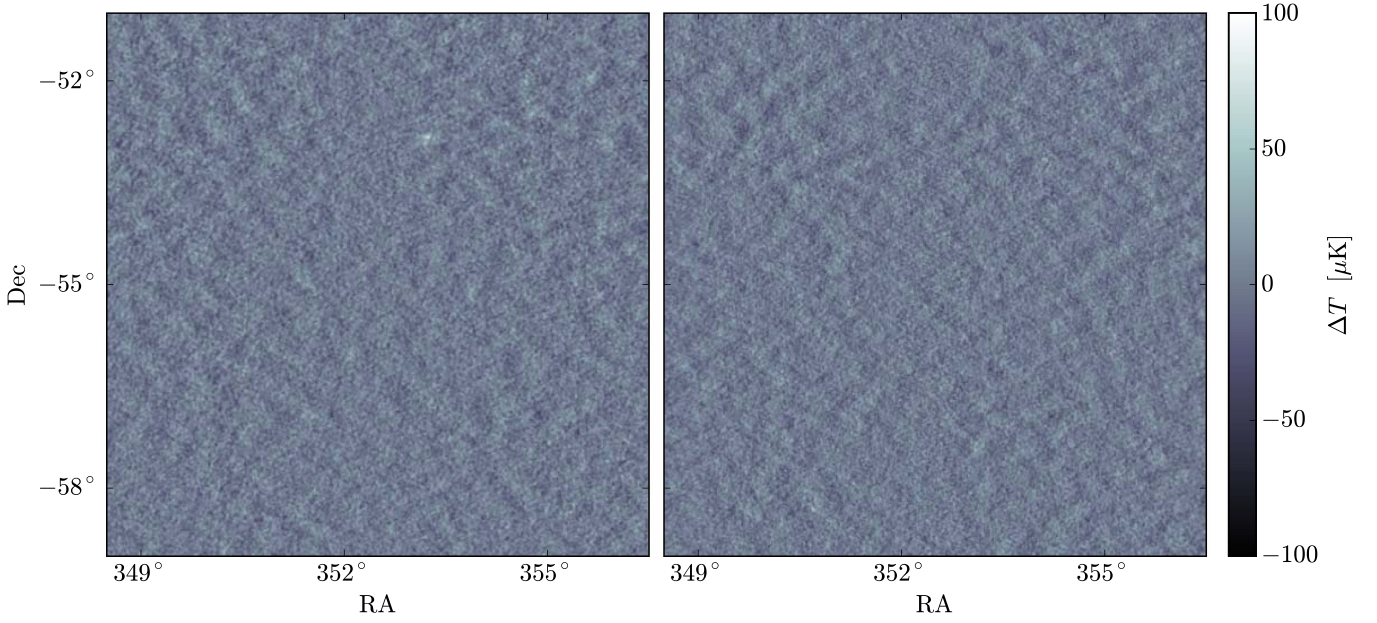


Figure 6. Left panel: residual map between the smoothed SPT-SZ 150 GHz map in the bottom left panel of Figure 5 and the high-pass-filtered *Planck* 143 GHz map in the bottom right panel of Figure 5. The feature at the location of one of the bright point sources is potentially due to temporal variability of the source, as the SPT and *Planck* data were not taken simultaneously. Right panel: same as the left panel, but with simulated data, consisting of the sum of an SPT-SZ noise realization smoothed to match *Planck* resolution and a *Planck* noise realization filtered to match the time-domain filtering of the SPT-SZ data.

143×143 bandpowers have slightly smaller error bars in this ℓ range because the modes lost in the SPT TOD filtering process result in slightly increased sample variance). Noise variance begins to dominate the 143×143 error bars at $\ell \gtrsim 1700$ and the 150×143 error bars at $\ell \gtrsim 2000$, while the 150×150 error bars are sample variance dominated over the entire range plotted.

The remaining panels of Figure 7 show various comparisons among the three sets of bandpowers. Two comparison schemes are applied: differences and ratios. The former is good for diagnosing additive effects, while the latter is more sensitive to multiplicative systematics. In both schemes, $\mathcal{D}_b^{150 \times 150}$ is chosen as the fiducial bandpower set. These comparison plots and the χ^2 and the probability to exceed (PTE) statistics calculated below are thus testing the following set of null hypotheses: (1) two sets of bandpower residuals

$$\begin{aligned} \langle \Delta \mathcal{D}_{b,c} \rangle &\equiv \langle \mathcal{D}_b^{150 \times 143} - \mathcal{D}_b^{150 \times 150} \rangle = 0 \\ \langle \Delta \mathcal{D}_{b,a} \rangle &\equiv \langle \mathcal{D}_b^{143 \times 143} - \mathcal{D}_b^{150 \times 150} \rangle = 0 \end{aligned}$$

and (2) two sets of bandpower ratios

$$\begin{aligned} \langle \Delta \mathcal{D}_{b,c} / \mathcal{D}_b^{150 \times 150} \rangle &\equiv \langle \mathcal{D}_b^{150 \times 143} / \mathcal{D}_b^{150 \times 150} \rangle - 1 = 0 \\ \langle \Delta \mathcal{D}_{b,a} / \mathcal{D}_b^{150 \times 150} \rangle &\equiv \langle \mathcal{D}_b^{143 \times 143} / \mathcal{D}_b^{150 \times 150} \rangle - 1 = 0. \end{aligned}$$

The absolute calibration of the SPT data from George et al. (2015) has a statistical uncertainty of $\sim 1\%$ at 150 GHz. We expect the bandpower comparison to be significantly more precise than this, so before plotting the bandpower residuals and ratios and before testing the null hypotheses above, we apply a recalibration parameter r_c to bandpowers containing SPT data:

$$\Delta \mathcal{D}_{b,c} \equiv r_c \mathcal{D}_b^{150 \times 143} - r_c^2 \mathcal{D}_b^{150 \times 150} \quad (11)$$

$$\Delta \mathcal{D}_{b,a} \equiv \mathcal{D}_b^{143 \times 143} - r_c^2 \mathcal{D}_b^{150 \times 150}, \quad (12)$$

and similarly for the bandpower ratios. Because of the precision of the George et al. (2015) calibration, we expect r_c to be very close to 1. We then calculate and minimize χ^2 as a function of r_c :

$$\chi^2(r_c) = \Delta \mathbf{D}_b^T(r_c) \Xi^{-1} \Delta \mathbf{D}_b(r_c), \quad (13)$$

and similarly for the bandpower ratios, where $\Delta \mathbf{D}_b$ denotes the vector that includes both $\Delta \mathcal{D}_{b,c}$ and $\Delta \mathcal{D}_{b,a}$. We fit the two sets of residuals or ratios simultaneously. In principle, we should include the recalibration parameter in an adjustment to the noise contribution to the covariance matrix, but we neglect it for simplicity, with the justification that the correction is very small at less than 1%. In Figure 7, we have included the best-fit recalibration parameter for SPT, r_c , in all three panels.

In the middle panel of Figure 7 we show the bandpower residuals $\Delta \mathcal{D}_{b,c}$ and $\Delta \mathcal{D}_{b,a}$ with error bars given by the square root of the diagonal elements of the full covariance matrix. This figure shows that the residual bandpowers are consistent with zero given the errors. In the bottom panels of the same figure we show the bandpower ratios. Similar to the residuals, the error bars of the bandpower ratios are the square root of the diagonal elements of the full covariance matrix. Qualitatively, these results appear to be consistent with our null hypotheses.

The quantitative characterization comes from the χ^2 value with the best-fit recalibration parameter. The best-fit values of r_c , minimum χ^2 , and probabilities to exceed that χ^2 are listed in Table 1 for several combinations of data and covariance. The primary results are from the combined residual bandpowers. Using the full 2×2 block covariance matrix for the recalibration fit, these results give a best fit of $r_c = 1.0087 \pm 0.0015$ with $\chi^2 = 78.7$. There are 37 ℓ bins in our analysis, so we have 74 data points among the two residual bandpowers and one free parameter. The PTE for $\chi^2 = 78.7$ and 73 degrees of freedom is 0.30. We find very similar results

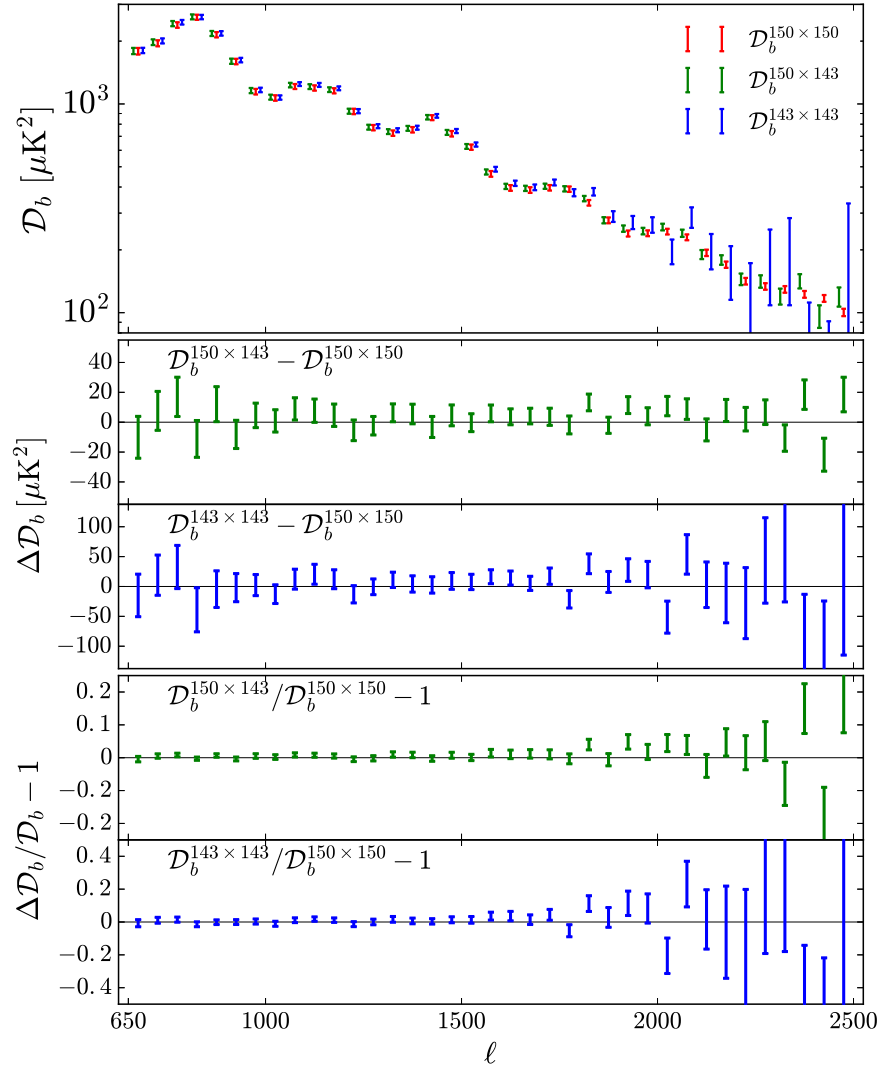


Figure 7. Top panel: window-function-corrected, unbiased bandpowers for 150×150 (red), 150×143 (green), and 143×143 (blue). The error bars contain sample variance and noise variance obtained from our simulations, as well as beam uncertainties from the analytical calculations presented in Section 3.4.2. The green and blue error bars are offset horizontally for clarity. Middle panel: bandpower residuals $\mathcal{D}_b^{150 \times 143} - \mathcal{D}_b^{150 \times 150}$ (green) and $\mathcal{D}_b^{143 \times 143} - \mathcal{D}_b^{150 \times 150}$ (blue) with the window function correction. The error bars come from the diagonal of the covariance matrix of the bandpower residuals of Equation (9). Under our null hypothesis, the residuals are expected to be consistent with 0. Note the different plotting range between the two bandpower residuals. Bottom panel: window-function-corrected bandpower ratios $\mathcal{D}_b^{150 \times 143} / \mathcal{D}_b^{150 \times 150} - 1$ (green) and $\mathcal{D}_b^{143 \times 143} / \mathcal{D}_b^{150 \times 150} - 1$ (blue). The best-fit recalibration parameter for SPT, r_c , has been applied to all panels of this figure.

from the combined-ratio fit: $r_c = 1.0092 \pm 0.0015$, $\chi^2 = 81.02$, PTE = 0.24. Put another way, given the noise properties and the beam uncertainties of the two experiments, 30% (24%) of our simulations have a higher χ^2 for the bandpower differences (ratios) than we find with the real data. We thus find as our primary result that the SPT and *Planck* data in the SPT-SZ sky patch are quite consistent with the null hypothesis that there is no systematic offset in the two experiments' measurement of the sky. A by-product of this analysis is that we recalibrate the SPT data with a statistical precision of 0.30% in power (0.15% in temperature) relative to *Planck* 143 GHz. This is comparable with the absolute calibration uncertainty of the *Planck* 143 GHz, 0.14% in power (0.07% in temperature; Planck Collaboration et al. 2016a), so we add this uncertainty in quadrature for a final SPT 150 GHz calibration uncertainty of 0.33% in power.

In Table 1, we also report the quantities of consistency from the single pair of bandpowers. For example, the PTE is 0.22 for

the residual between $\mathcal{D}_b^{150 \times 143}$ and $\mathcal{D}_b^{150 \times 150}$ with the full covariance matrix. For the other pair of residual bandpowers, $\mathcal{D}_b^{143 \times 143}$ and $\mathcal{D}_b^{150 \times 150}$, the PTE is 0.43 with the full covariance matrix. The results without the beam uncertainties are also listed in Table 1.

As discussed in Section 3.2.2, the white-noise assumption in simulations for $\mathcal{D}_b^{143 \times 143}$ results in an overestimate of the noise contribution to the covariance matrix, at roughly the 10% level in \mathcal{D}_b errors, or the 20% level in variance. If we assumed that the *Planck* noise was the dominant contribution to the residual and ratio covariance in the 150×150 versus 143×143 comparisons, we would expect roughly a 20% increase in χ^2 if we were able to use more realistic noise simulations. The resulting χ^2 would still correspond to a reasonable PTE ($\gtrsim 0.16$) for the null hypothesis. This would also be true of the combined constraints, particularly because the main constraining power comes from $\mathcal{D}_b^{150 \times 143}$, which is unaffected by the assumption of white noise in the *Planck*-only bandpowers.

Table 1
Best-fit Recalibration Parameter, χ^2 , and PTE for the Various Bandpower Comparisons

Comparison	Best r_c	χ^2	PTE
<i>Residual, combined, full covariance</i>	1.0087 ± 0.0015	78.7	0.30
Residual, combined, no beam error	1.0117 ± 0.0009	89.8	0.09
150×143 – 150×150 , full covariance	1.0087 ± 0.0016	42.1	0.22
150×143 – 150×150 , no beam error	1.0115 ± 0.0010	50.4	0.06
143×143 – 150×150 , full covariance	1.0076 ± 0.0022	36.8	0.43
143×143 – 150×150 , no beam error	1.0110 ± 0.0013	41.8	0.23
<i>Ratio, combined, full covariance</i>	1.0092 ± 0.0015	81.0	0.24
Ratio, combined, no beam error	1.0120 ± 0.0009	91.0	0.08
$(150 \times 143)/(150 \times 150)$, full covariance	1.0090 ± 0.0016	43.3	0.19
$(150 \times 143)/(150 \times 150)$, no beam error	1.0118 ± 0.0010	50.9	0.05
$(143 \times 143)/(150 \times 150)$, full covariance	1.0082 ± 0.0022	38.1	0.37
$(143 \times 143)/(150 \times 150)$, no beam error	1.0113 ± 0.0013	42.4	0.21

Because of the largely reduced sample variance contribution to the difference between bandpowers within the same sky coverage (roughly a factor of six in $150 \times 143 - 150 \times 150$ compared to 150×150), the comparison of $\mathcal{D}_b^{150 \times 150}$ with the $\mathcal{D}_b^{150 \times 143}$ and $\mathcal{D}_b^{143 \times 143}$ bandpowers derived from within the SPT patch provides tighter tests, over a wide range of angular scales, than can be achieved by the comparison of $\mathcal{D}_b^{150 \times 150}$ with the more precise *Planck* spectra derived from nearly the full sky. In Figure 8, we compare the uncertainties on the two sets of bandpower residuals $\mathcal{D}_b^{143 \times 143} - \mathcal{D}_b^{150 \times 150}$ (blue) and $\mathcal{D}_b^{150 \times 143} - \mathcal{D}_b^{150 \times 150}$ (green) to the uncertainties on the SPT-only bandpowers $\mathcal{D}_b^{150 \times 150}$ (red). In the lower- ℓ region ($\ell < 1800$), the green curve has the lowest error because the sample variance has been greatly reduced by subtracting the SPT bandpowers from the cross-bandpowers, while at higher ℓ the green curve rises owing to the *Planck* noise contribution. In the most constraining case ($\mathcal{D}_b^{150 \times 143} - \mathcal{D}_b^{150 \times 150}$), the errors for almost all bins are $\leq 10 \mu\text{K}^2$, comparable to the uncertainty in the PlanckFS bandpowers. Both sets of bandpower residuals yield very stringent tests on the consistency of the two data sets, and our results show that these data sets are formally consistent in this patch.

4.3. Comparison to the Full-sky Planck 2015 TT High- ℓ Bandpowers

While restricting the bandpower comparison to overlapping sky substantially reduces differential sample variance, it does increase the *Planck* covariance. In this section, we instead investigate differences between the “in-patch” bandpowers (i.e., the 150×150 , 150×143 , and 143×143 bandpowers from the 2540 deg^2 SPT-SZ survey region) and the full-sky *Planck* 2015 TT high- ℓ unbinned frequency combined bandpowers (Planck Collaboration et al. 2016b), which we refer to as “PlanckFS.” In this case, the null hypothesis is that any differences between PlanckFS and the in-patch bandpowers are consistent with expectations given Gaussian statistics and statistical isotropy.

In Figure 9 we show the residuals and ratios for the in-patch to PlanckFS bandpowers. The residual plots show that all three

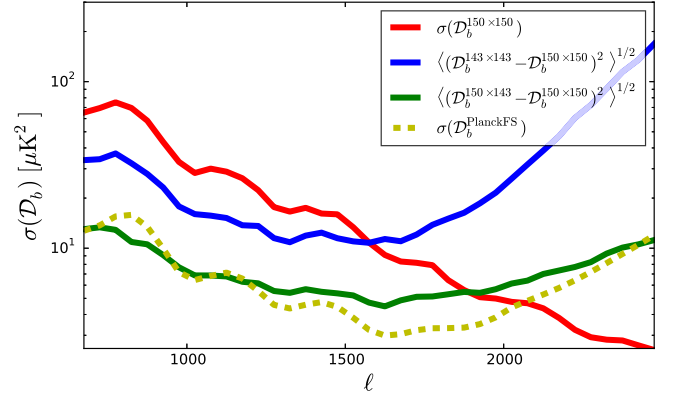


Figure 8. Uncertainties in the SPT 150×150 bandpowers (red) and in the two sets of bandpower residuals (not including beam uncertainty). The sample variance is significantly reduced in the $150 \times 150 - 143 \times 143$ residuals compared to the 150×150 bandpowers (by roughly a factor of six in the lowest ℓ ranges, and by a larger factor at higher ℓ), indicating that this comparison is more stringent than the comparison of 150×150 to full-sky *Planck* data, the uncertainty on which is dominated by the sample variance in the 150×150 spectrum. Also plotted are the bandpower uncertainties from full-sky *Planck* data, scaled to the ℓ bin size used in this work.

sets of the in-patch bandpowers prefer greater power at $650 < \ell < 1200$, indicating that the SPT-SZ patch has greater power at these multipoles than the full-sky average. The ratio plots also suggest a tilt with respect to PlanckFS, although the significance is modest. The slope of the tilt is consistent to 0.5σ between the three cross-spectra, being slightly larger in 150×150 and smaller in 143×143 .

To quantify the statistical significance of this tilt, we model the ratio of the in-patch and PlanckFS bandpowers as a power law:

$$\frac{D_\ell^{i \times j}}{D_\ell^{\text{PlanckFS}}} = A \left(\frac{\ell}{4000} \right)^n. \quad (14)$$

We assume a Gaussian likelihood with

$$-2 \ln \mathcal{L}(A, n | D^{i \times j}, D^{\text{PlanckFS}}) = \Delta_b \Sigma_{bb'}^{-1} \Delta_b', \quad (15)$$

$$\Delta_b = \frac{D_b^{i \times j}}{W_{bt}^{i \times j} D_\ell^{\text{PlanckFS}}} - R_b^{i \times j}(A, n), \quad (16)$$

$$R_b^{i \times j}(A, n) = W_{bt}^{i \times j} \left(A \left(\frac{\ell}{4000} \right)^n + F_\ell^{i \times j} \right). \quad (17)$$

Here D_ℓ^{PlanckFS} is the best-fit PlanckFS power spectrum, $\Sigma_{bb'} = \langle \Delta_b \Delta_{b'} \rangle$ (only the SPT beam uncertainty is included and cross-correlations between $D_b^{i \times j}$ and the D_ℓ^{PlanckFS} are negligible), and $F_\ell^{i \times j}$ is the foreground model adopted from S13 with the frequency dependence from George et al. (2015) included. There are three foreground amplitude parameters included in $F_\ell^{i \times j}$ to account for the SZ, Poisson-point-source, and clustered CIB uncertainties. The best-fit recalibration from Section 4.3 has been applied to the in-patch SPT-only and SPT-*Planck* bandpowers. The best-fit values for A and n are reported in Figure 9. We find only marginal evidence for a tilt between PlanckFS and any of the three cross-spectra; the most significant tilt is for 150×150 and is 1.5σ away from zero. We conclude that the tilts we see in the observed spectra are roughly consistent with the expectation based on Gaussian

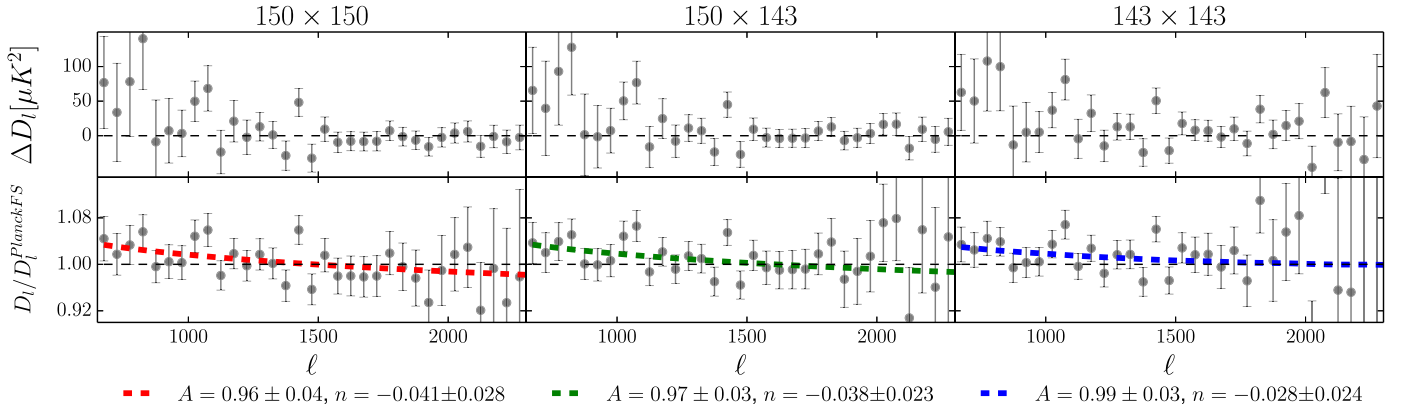


Figure 9. Top panel: residuals between the in-patch and PlanckFS bandpowers. Bottom panel: ratios of the in-patch and PlanckFS bandpowers. The colored dashed lines are the best-fit power laws for each ratio.

statistics and the assumption of statistical isotropy—i.e., they are roughly consistent with $n = 0$.

To determine the expected statistical properties of the differences in the best-fit values of A and n , we construct a covariance matrix (\mathbf{C}) from the set of best-fit values of A and n in each of the 400 simulations of Section 3.2. We then calculate χ^2 for these differences as

$$\chi^2 = (\boldsymbol{\Theta} - \bar{\boldsymbol{\Theta}})^T \mathbf{C}^{-1} (\boldsymbol{\Theta} - \bar{\boldsymbol{\Theta}}), \quad (18)$$

where $\boldsymbol{\Theta}$ is the vector of parameter differences and $\bar{\boldsymbol{\Theta}}$ is the mean simulation difference, which is consistent with zero. The breakdown of PTEs for the various pair differences is shown in Table 2. As can be seen, the most extreme PTE is 0.85 and the lowest is 0.41. We conclude that the observed tilts in the three cross-spectra are completely consistent with each other.

All our tests are consistent with the following explanation for the tilts we observe in the in-patch spectra relative to the best-fit *Planck* full-sky spectrum: they are driven by a sample variance fluctuation away from the full-sky average, the magnitude of which is roughly consistent with expectations under the assumption of statistical isotropy and our noise model.

4.4. Pipeline Checks

All published power spectra from the SPT Collaboration have been calculated using some variant of the cross-spectrum pseudo- C_ℓ pipeline used in this work. Extensive checks have been performed on this pipeline (see, e.g., Section 4.2 of Story et al. 2013), demonstrating that the correct input spectrum is recovered from simulated data, even when that spectrum differs from the spectrum assumed in calculating the filter transfer function. If, however, some aspect of the pipeline were inducing a bias in the estimated power spectra (through some mechanism that has escaped all pipeline tests), this bias would affect both the SPT-SZ and *Planck* data used in this work (because we have mock-observed the *Planck* data and analyzed it with the SPT pipeline). If the bias on the two data sets were comparable, it would then divide or subtract out in the bandpower comparison, and we would (wrongly) conclude that there was no issue with either data set.

To test this scenario, we have analyzed the in-patch *Planck* data using an alternate pipeline. Specifically, we created a HEALPix version of the SPT-SZ sky-patch and point-source

Table 2
Power-law Parameter PTEs

	χ^2	PTE
150 × 150 – 150 × 143	0.33	0.41
150 × 150 – 143 × 143	1.79	0.85
150 × 143 – 143 × 143	1.18	0.55

Note. The χ^2 and PTE for the null hypothesis that the tilt relative to PlanckFS is the same in the in-patch 150 × 150, 150 × 143, and 143 × 143 bandpowers.

mask (stitched together from the individual-field masks), and we handed this mask and the half-mission *Planck* 143 GHz maps to the PolSpice⁴⁰ (Szapudi et al. 2001; Chon et al. 2004) code, which is designed to estimate the power spectra of masked full-sky maps and properly account for the masking. We binned the $\Delta\ell = 1$ PolSpice output into the $\Delta\ell = 50$ bins used in the SPT pipeline using the bandpower window functions calculated for the 143 × 143 “scanned, filtered” *Planck* bandpowers. We then calculated the ratio of the “unscanned, unfiltered” (PolSpice) bandpowers to the scanned, filtered ones and found that they agree to better than 3% in every individual bin in which there is appreciable signal-to-noise ratio in the 143 × 143 spectrum, with an overall ratio of 1.0028 ± 0.0050 over the range $600 < \ell < 1800$ and no evidence of a trend with ℓ . We have also redone the tilt calculation in Section 4.3 using the unscanned, unfiltered bandpowers and found results consistent with what we found with the scanned, filtered bandpowers (within a fraction of a sigma). We thus conclude that our fundamental results are not an artifact of the SPT analysis method.

5. Conclusions

In this paper, we have compared 150 GHz SPT data and 143 GHz *Planck* data in the same region of the sky, namely, the 2540 deg² SPT-SZ survey footprint. We have performed a visual comparison of maps constructed from the two data sets and found the difference between the two maps to be visually consistent with noise, once they have been filtered to display the same angular modes.

We then performed a quantitative analysis of the consistency of the maps, relying primarily on a comparison of the

⁴⁰ <http://www2.iap.fr/users/hivon/software/PolSpice>

cross-spectrum of two halves of the SPT data with the $\text{SPT} \times \text{Planck}$ cross-spectrum. We also compared the $\text{SPT} \times \text{SPT}$ spectrum with the cross-spectrum of two halves of the *Planck* data. These comparisons were made using differences between and ratios of two sets of binned power spectra, or bandpowers, at a time, always using the $\text{SPT} \times \text{SPT}$ bandpowers as the fiducial set. To test the null hypothesis that the bandpower differences (after recalibrating the SPT data) are consistent with zero—or that the ratios are consistent with unity—we created a suite of 400 simulations of the signal and noise properties of the SPT and *Planck* maps, including signal contributions from the CMB and extragalactic foregrounds. We found our most stringent test, based on the expected variance of the differences, to be the comparison of the $\text{SPT} \times \text{SPT}$ spectrum with the SPT-Planck cross-spectrum. Forming a χ^2 quantity from these bandpower differences and the simulation-based covariance matrix, we have found a value that is exceeded by 22% of the analogous χ^2 values derived from the simulated data, i.e., corresponding to the PTE value 0.22. When we add the residuals between the $\text{SPT} \times \text{SPT}$ and $\text{Planck} \times \text{Planck}$ cross-spectra, we find a PTE of 30%. All other tests result in similarly unremarkable PTEs. We find no evidence of a failure of our null model, i.e., we see no evidence for systematic errors or under- or overestimates of statistical errors.

We have also compared the three sets of bandpowers from the 2540 deg^2 SPT-SZ survey region to the full-sky *Planck* 143 GHz power spectrum. Relative to the *Planck* full-sky spectrum, we have found a hint for a tilt in the in-patch bandpowers. For all three sets of in-patch bandpowers, the amplitudes of the tilts we have obtained are consistent with each other and roughly consistent with expected noise and sample variance fluctuations.

This work shows that the SPT 150 GHz and *Planck* 143 GHz data are in very good agreement with each other within the 2540 deg^2 SPT-SZ survey area. In a companion paper (Aylor et al. 2017), we extend this comparison to the cosmological parameters that can be derived from these bandpowers.

The South Pole Telescope is supported by the National Science Foundation through grant PLR-1248097. Partial support is also provided by the NSF Physics Frontier Center grant PHY-1125897 to the Kavli Institute of Cosmological Physics at the University of Chicago, the Kavli Foundation, and the Gordon and Betty Moore Foundation grant GBMF 947. The McGill group acknowledges funding from the National Sciences and Engineering Research Council of Canada, the Canada Research Chairs program, and the Canadian Institute for Advanced Research. Argonne National Laboratory work was supported under U.S. Department of Energy contract DE-AC02-06CH11357. B.B. is supported by the Fermi Research Alliance, LLC under contract No. DE-AC02-07CH11359 with the U.S. Department of Energy. C.R. acknowledges support from an Australian Research Councils Future Fellowship (FT150100074). We thank D. Rapetti for comments on the manuscript, and we thank an anonymous referee for helpful suggestions.

Appendix



In this appendix, we explain why we use a different power spectrum estimator for the SPT bandpowers than in S13, and

we discuss the potential difference this makes to our comparisons. S13 averaged the cross-spectra between $\mathcal{O}(100)$ single-observation maps, while in this work we use the cross-spectrum between two half-survey maps in each field. Similarly, S13 estimated the bandpower covariance matrix from the distribution of said cross-spectra, while we estimate the covariance from signal+noise simulations.

The decision to change estimators is driven by the desire to use the same procedure for the 150×150 , 150×143 , and 143×143 bandpowers. With only the two half-survey maps for 143 GHz, the S13 covariance estimator would not work. Instead, we rely on signal+noise simulations to calculate the bandpower covariance. However, we could only create one noise realization per map for the single-observation maps used in S13. By using half-survey maps (i.e., by co-adding many single-observation maps), we can increase the number of independent noise realizations dramatically. As laid out in Section 3.2.3, we generate noise realizations by first nulling the signal in single-observation maps by differencing the left-going and right-going scans and then co-adding these noise maps with a random $+1$ or -1 prefactor. These noise realizations can then be added to the simulated signal-only maps to yield robust signal+noise SPT map simulations.

One might worry that the change in estimator could affect our comparison. To address this, we perform the following quantitative test. We replace the SPT 150 GHz half1-half2 cross bandpower by the original S13 bandpowers and redo the test on the bandpower residuals (with the covariance matrix of the residual unchanged). We find no significant differences. For $\mathcal{D}_b^{150 \times 143 - 150 \times 150}$ only with the S13 bandpowers, the best-fit $r_c = 1.0090 \pm 0.0016$ with $\chi^2 = 47.21$ and PTE = 0.100. Recall from Table 1 that the numbers with the half-survey bandpowers are $r_c = 1.0087 \pm 0.0016$ with $\chi^2 = 42.07$ and PTE = 0.224. For $\mathcal{D}_b^{143 \times 143 - 150 \times 150}$ with the S13 bandpowers, the best-fit $r_c = 1.0086 \pm 0.0022$ with $\chi^2 = 39.17$ and PTE = 0.330. Again from Table 1, the equivalent numbers with the half-survey bandpowers are $r_c = 1.0076 \pm 0.0022$ with $\chi^2 = 36.83$ and PTE = 0.430. We therefore conclude that the change in SPT power spectrum estimators does not significantly impact our comparison of the SPT 150 GHz map and *Planck* 143 GHz map over the same region of sky.

ORCID iDs

T. M. Crawford  <https://orcid.org/0000-0001-9000-5013>
T. de Haan  <https://orcid.org/0000-0001-5105-9473>
E. M. George  <https://orcid.org/0000-0001-7874-0445>
G. P. Holder  <https://orcid.org/0000-0002-0463-6394>
D. P. Marrone  <https://orcid.org/0000-0002-2367-1080>
Y. Omori  <https://orcid.org/0000-0002-0963-7310>
C. L. Reichardt  <https://orcid.org/0000-0003-2226-9169>
A. A. Stark  <https://orcid.org/0000-0002-2718-9996>
J. D. Vieira  <https://orcid.org/0000-0001-7192-3871>

References

- Addison, G. E., Huang, Y., Watts, D. J., et al. 2016, *ApJ*, **818**, 132
- Aylor, K., Hou, Z., Knox, K. T., et al. 2017, *ApJ*, **850**, 101
- Battye, R. A., & Moss, A. 2014, *PhRvL*, **112**, 051303
- Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, *ApJS*, **208**, 20
- Cai, Z.-Y., Lapi, A., Xia, J.-Q., et al. 2013, *ApJ*, **768**, 21
- Calabrese, E., Hlozek, R. A., Battaglia, N., et al. 2013, *PhRvD*, **87**, 103012
- Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, *MNRAS*, **350**, 914

- De Zotti, G., Ricci, R., Mesa, D., et al. 2005, [A&A](#), **431**, 893
- Efstathiou, A., Pearson, C., Farrah, D., et al. 2014, [MNRAS](#), **437**, L16
- George, E. M., Reichardt, C. L., Aird, K. A., et al. 2015, [ApJ](#), **799**, 177
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, [ApJ](#), **622**, 759
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, [ApJS](#), **208**, 19
- Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, [ApJ](#), **567**, 2
- Hou, Z., Reichardt, C. L., Story, K. T., et al. 2014, [ApJ](#), **782**, 74
- Keisler, R., Reichardt, C. L., Aird, K. A., et al. 2011, [ApJ](#), **743**, 28
- Knox, L. 1999, [PhRvD](#), **60**, 103516
- Lewis, A. 2005, [PhRvD](#), **71**, 083008
- Louis, T., Addison, G. E., Hasselfield, M., et al. 2014, [JCAP](#), **7**, 016
- Mocanu, L. M., et al. 2013, [ApJ](#), **779**, 61
- Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2016a, [A&A](#), **594**, A8
- Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2016f, [A&A](#), **594**, A7
- Planck Collaboration, Ade, P. A. R., & Aghanim, N. 2016d, [A&A](#), **594**, A12
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014a, [A&A](#), **571**, A1
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014b, [A&A](#), **571**, A7
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014c, [A&A](#), **571**, A16
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, [A&A](#), **594**, A13
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016c, [A&A](#), **594**, A24
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2017, [A&A](#), **607**, A95
- Planck Collaboration, Aghanim, N., Arnaud, M., et al. 2016e, [A&A](#), **594**, A11
- Riess, A. G., Macri, L., Casertano, S., et al. 2011, [ApJ](#), **730**, 119
- Riess, A. G., Macri, L., Hoffmann, S. L., et al. 2016, [ApJ](#), **826**, 56
- Schaffer, K. K., Crawford, T. M., Aird, K. A., et al. 2011, [ApJ](#), **743**, 90
- Shaw, L. D., Nagai, D., Bhattacharya, S., & Lau, E. T. 2010, [ApJ](#), **725**, 1452
- Shaw, L. D., Rudd, D. H., & Nagai, D. 2012, [ApJ](#), **756**, 15
- Snyder, J. P. 1987, Map Projections-a Working Manual (Washington, DC: Geological Survey)
- Story, K. T., Reichardt, C. L., Hou, Z., et al. 2013, [ApJ](#), **779**, 86
- Szapudi, I., Prunet, S., Pogossyan, D., Szalay, A. S., & Bond, J. R. 2001, [ApJL](#), **548**, L115
- Wyman, M., Rudd, D. H., Vanderveld, R. A., & Hu, W. 2014, [PhRvL](#), **112**, 051302